CLAIMS

1. A method for searching for an input symbol string among a set of symbol strings, comprising:

creating (A) a trie data structure of symbol strings, wherein the symbol strings are grouped into branches in such a manner that the symbol strings beginning with the same symbols belong to the same branch, and the symbol strings in the same branch divide into new branches at the symbols, from which onwards the symbols strings differ from each other,

receiving (B) an input formed of an input symbol string,

proceeding (C) from the starting point of the trie data structure along a branch to a calculation point indicated by the next symbol,

calculating (D) distances at the calculation point between a sample symbol string formed by the symbols of the calculation point of the branch in question and the calculation points preceding it and the input symbol string by comparing these in alternative ways,

selecting (E) repeatedly the next branch to follow (C) to the calculation point indicated by the next symbol, at which said calculation (D) is repeated for the new calculation point,

after the calculation (G) has terminated, selecting one or more symbol strings having the shortest distance to the input symbol string on the basis of the performed calculations, and

using the selected symbol string(s) to produce a response, **characterized** by

calculating (D) at the calculation points not only the distances, but also the smallest possible length difference corresponding to each distance that indicates how much the length of the remaining part of the input symbol string not examined in the distance calculation differs from the lengths remaining in the symbols strings passing through the calculation point, and calculating on the basis of each distance and corresponding length difference a reference value, and

performing (E) said selection of the next branch in such a manner that next the routine proceeds from the calculation point, from which the lowest reference value has been obtained as result.

2. A method as claimed in claim 1, **characterized** by

comparing the distance of the symbol string or strings used to produce the response and that of the input symbol string with a predefined maximum distance, and

changing the produced response to indicate that the input symbol string was not found if the distance exceeds the maximum distance.

3. A method as claimed in claim 1 or 2, **characterized** by

when selecting the branch, comparing said lowest reference value with the predefined maximum distance, and

terminating the calculation if the lowest reference value exceeds the maximum distance.

4. A method as claimed in any one of claims 1 to 3, **characterized** by

when selecting the branch, checking whether calculation is already done for the last calculation point on one of the branches, and

terminating the calculation, if it turns out that for the last calculation point of one of the branches a reference value has been obtained that is lower than the reference values obtained for all the other calculation points.

5. A computer program for implementing the method according to any one of claims 1 to 4.

6. A data medium readable by computer, **characterized** in that a computer program is readable from the data medium for performing the method according to any one of claims 1 to 4.

7. An apparatus (10) for searching for a symbol string among a set of symbol strings, the apparatus comprising:

means (12) for creating a trie data structure of symbols strings by grouping the symbol strings into branches in such a manner that the symbol strings starting with the same symbols belong to the same branch, and the symbols strings in the same branch divide into new branches at the symbols, from which onwards the symbol strings differ from each other,

an input (11) for receiving an input symbol string,

calculation means (14) for calculating distances between a sample symbol string formed by the symbols of the calculation point and the calculation points preceding it in the examined branch and the input symbol string by comparing these in alternative ways,

selection means (15) that repeatedly select the next branch to follow to the calculation point indicated by the next symbol, at which said calculation is repeated for the new calculation point,

selection means (16) that, after the calculation is terminated, select one or more symbol strings with the shortest distance to the input on the basis of the calculations,

response production means (17) that produce a response by using the selected symbol string(s), and

an output (18) for feeding the response onward, **characterized** in that the apparatus is arranged to

calculate at the calculation points not only the distances, but also the smallest possible length difference corresponding to each distance that indicates how much the length of the remaining part of the input symbol string not examined in the distance calculation differs from the lengths remaining in the symbols strings passing through the calculation point, and calculating on the basis of each distance and corresponding length difference a reference value, and

perform said branch selection in such a manner that next the routine proceeds from the calculation point, from which the lowest reference value has been obtained as result.